

# Review Articles

## Trends in Documentation

*Modern Trends in Documentation; Proceedings of a Symposium held at the University of Southern California, April 1958.* Ed. by Martha Boaz. New York: Pergamon Press, 1959. viii, 103p. \$5.00.

The symposium reported in this modest volume represents an attempt by the School of Library Science of the University of Southern California to stimulate interest in information retrieval, which is one aspect of documentation. The program consisted of three general presentations on the planning, nature, and needs of information retrieval; two reports on machine translation; three outlines of information retrieval systems utilizing computers; two descriptions of electronic devices and their uses; and, finally, a short panel discussion.

The symposium opened with a discussion by Robert Meyer of the needs of the user of information, largely in terms of the user in the special library. This was followed by papers of H. J. R. Grosch (IBM) and Merritt Kastens (Stanford Research Institute), clearly setting forth the objectives of information retrieval. All three presentations were very explicit as to the type of information needed by the scientific and technical specialist, including explanation of the kind of occasion calling for speed, which has been one of the primary factors in attempts at mechanization of the information-finding processes. It should be emphasized that the approach to knowledge represented in the machine methods discussed by these speakers was essentially a subject one.

The machine translation section of the conference, covered by H. P. Edmundson and D. G. Hays of the Rand Corporation, formed a very lucid introduction to the subject. Those who attended the symposium must have been well prepared to tackle the papers on mechanical translation in the preprints sent out for the International Conference on Scientific Information held seven months later. All of the terms used to describe the machine translation procedures have been explained for the layman.

During the symposium, three information retrieval systems were described by H. P. Luhn (IBM), Don Andrews (U. S. Patent Office), and Harley Tillett (U. S. Naval Ordnance Test Station, China Lake, Calif.). These showed how machines designed for mathematical purposes could be utilized to find certain types of information. In all three cases the organization of knowledge was rearranged to fit the computer. Luhn's "auto-encoding" of documents is an interesting system. "Auto-encoding" is a subject-entry method based on the choice of key words whose importance is determined by their statistical frequency in the documents analyzed. The weakness of the method for broader application lies in the fact that many authors, for aesthetic reasons, deliberately try *not* to use the same major subject word twice in close proximity, preferring as many synonyms as feasible to avoid annoying the reader with constant repetition. Without a very fancy thesaurus, this would tend to spoil the statistical averages and make all key words of equal value. Tillett's paper mentioned some psychological factors affecting the adoption of machine methods for information searching. These factors, notably disappointment that the machine was so slow and that it could not do everything for the client, could be overcome by educating the user not to ask a machine a question that could be answered better with a dictionary catalog, or, in this case, the Uniterms system; and also by constantly reminding the client that any machine is a mechanical moron created to perform tedious or repetitious tasks, but not to do anything requiring much intelligence.

No information systems were described of the type which utilize machines specifically for bibliographic rather than mathematical or statistical purposes, such as semantic factoring and the Western Reserve Searching Selector. However, general features of several machines for information retrieval were presented by Peter Worsley (Benson-Lehner), who described a kind of Rapid Selector called FLIP, and by Robert Hayes (Magnavox), who described Minicards and Magnacards, both of which are entirely different

in design and operation from computers.

The panel discussion at the end of the Symposium revealed the extent to which its objectives were achieved or missed. Apparently the term "documentation" was not defined at the beginning, for one finds reference among the comments of the panelists (p. 89) to an article of Harry Bauer in which documentation has been described as a means of organizing files of letters and pamphlet-like material, a misconception almost as common as the one which defines documentation primarily in terms of the care and feeding of data-processing machines. The field of documentation is much broader than either of these views. A "document" is any medium containing recorded evidence of intellectual endeavor. "Documentation" is used in the *narrow* sense by historians to mean the process of citing written evidence to substantiate a statement of fact and also as the name for such evidence. The word is used in the *broad* sense by those working with collections of data, or information of any kind, to mean any process connected with the "identification, recording, organization, storage, recall, conversion into more useful forms, synthesis, and dissemination of the *intellectual content* of print or any other recorded materials."<sup>1</sup> The dual meaning of the word "documentation" is quite clear if one remembers that a historian will accept the fact that a decisive battle was fought at Hastings in 1066, but will insist on "documenting" a statement that victory in this battle was influenced by the ability of heavily armored Norman knights to fight on horseback, because they rode a new breed of horses large enough to support a knight in full armor, while the English rode to the battle field on small horses, then had to dismount and fight on foot. The documentalist, on the other hand, will treat battle, place, date, English knights, Norman knights, infantry, cavalry, full armor, light armor, big horses, and little horses impartially as "information" to be processed, stored, and recovered. The reasoning of the historian, the thesis for which he had to cite chapter and verse, will only find its way into this body of information through the further addition of some generalized subject headings, such as "tactics," "military art and

science," "military history, medieval," or "armies—equipment." Incidentally, this example points up another problem in Luhn's "auto-encoding": the significant ideas or conclusions in a document are not necessarily expressed in precise terms suitable for storage and retrieval, while the wordy argument used in establishing them may not be worth preserving. (The example given here is the late Carl Stephenson's *Big Horse Theory of the Battle of Hastings*.)

If the meaning of "documentation" was not made clear to the librarians on the panel of the symposium, it is also obvious that the necessities of library work were equally vague to the machine and system makers present. At one point (pp. 96-97) it was suggested that it is the job of the "mother" professional organization in the library world (presumably ALA) to formulate a set of standards for mechanization. This idea seems to crop up in one form or another at every documentation conference. It might have a chance of being adopted if there were somewhere a plain statement of exactly what the machines can and cannot do similar to the beginning made by Claire K. Schultz at the recent meeting of the American Documentation Institute. It also might be acted upon if there were some indication of interest in producing a kind of machine that would fit library procedures, rather than demanding that these experience-tested processes be turned upside down to fit a machine designed for a counting-house. Actually, there are some routines which could be mechanized right now. For example, the practicing cataloger who does *original* cataloging could use something in the line of mechanical subject indexing to do these things: (1) take the raw terms which describe the subject matter of the book in hand, compare these with the standardized list of subject headings, and convert raw terms into the equivalent standardized forms; (2) take the title (or subtitle or supplied title) of the book as representative of its topic, together with the standardized subject heading most nearly descriptive of this topic, and compare with other titles already listed under the same heading to see whether the book being cataloged fits the category (if not, the cataloger may either try other subject headings turned up in step (1) or repeat the first process with further

<sup>1</sup> Ralph R. Shaw, "Documentation: Complete Cycle of Information Service," *CRL*, XVIII (1957), 452. (Italics are the reviewer's).

instructions refining the search among standardized headings); (3) repeat the same comparison and checking procedures with the classification tables. It might be possible to do both the subject heading and classification comparisons simultaneously. Present machines are capable of performing this kind of look-up and comparison operations.

During the symposium, the information needs of the scientific world, at least, were very clearly described, and a few of the imaginative methods which have been made by scientists and engineers to answer these needs were outlined. It is doubtful that the conference did much to end the Great Schism between the librarians, who understand the magnitude of the information storage and retrieval problem in its totality, and the proponents of mechanization, who see only the failures of present systems in the highly specialized fields with which they are familiar. Dean Boaz and the library school of the University of Southern California are to be commended on making a sincere effort to open channels of communication between the two viewpoints.—Phyllis A. Richmond, *University of Rochester Library*.

## Audio-Visual Tool

*They See What You Mean.* By [Eric F. Bur-  
tis and James E. LeMay] Ozalid Audio-  
Visual Department. Johnson City, N. Y.:  
Ozalid Division of General Aniline and  
Film Corporation, 1959. 88p., \$3.75.

The overhead projector is an audio-visual tool that has appeared since 1950 and made its presence felt quite markedly in industrial audio-visual departments. It is beginning to appear at technical and academic meetings as an extension to the services offered by the older projection methods. It combines the freedom and spontaneity of the blackboard with the precision and artistry of the slide projector, while adding a number of facilities not found in these standard techniques.

Ozalid does not make overhead projectors, but it does produce equipment and supplies used in preparing transparencies for these projectors. The bulk of this superbly illustrated volume deals with the preparation of transparencies by the diazo process. This is

to be expected as Ozalid is the outstanding producer of diazo materials in this country. Sections of the book are given to homemade transparencies, Transferon (diffusion-transfer), transparency mounting techniques, transparency design, and overhead projection techniques. Much of this manual could be used to improve presentations based on the blackboard and slide projector, and it will certainly add to the versatility of the department using an overhead projector.

There is a short bibliography at the end of the book referring the reader to sixteen recent reports on overhead projection. There is also a two-page listing of Ozalid audio-visual products, which serves as a glossary to the many terms savoring of jargon that appear in the book. The illustrations with which the book is filled serve to simplify the description of techniques and exemplify the visual method at its best. The profusion of trade-names in the text tends to minimize the effectiveness of this portion of the book. The volume can be recommended for all libraries engaged in or about to become involved in audio-visual work.—Hubbard W. Ballou, *Columbia University Libraries*.

## Soviet Publishing

*Publishing in the U.S.S.R.* By Boris I. Gorokhoff. (Indiana University Publications. Slavic and East European Series, Vol. 19.) [Bloomington, Ind.: Indiana University, c1959.] xvi, 307 p. \$3.00; cloth, \$6.00.

In spite of the recent burgeoning of articles about Soviet methods of disseminating scientific information there has been a need for full length studies in English which would give a balanced presentation of Soviet libraries, bibliography, and publishing in general. The Council on Library Resources, Inc., has acted to fill this need by supporting Paul Horecky's *Libraries and Bibliographic Centers in the Soviet Union*, Volume 16 in the Indiana series, and its companion volume on publishing. Together they form a valuable survey of the current scene. The competence in research on Soviet Russia, built up in large measure since World War II by the area institutes in American universities, appears to have been joined happily with experienced librarianship in the production of these studies. A volume on